EXHIBIT 8

UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

AUTHORS GUILD, *et al.*, individually and on behalf of others similarly situated,
 	
   Plaintiffs,
 	
     v.
 	
OPENAI INC., OPENAI OPCO LLC, OPENAI GP LLC, OPENAI LLC, OPENAI GLOBAL LLC, OAI CORPORATION LLC, OPENAI HOLDINGS LLC, OPENAI STARTUP FUND I LP, OPENAI STARTUP FUND GP I LLC, OPENAI STARTUP FUND MANAGEMENT LLC, and MICROSOFT CORPORATION,
 	
   Defendants.
	Case No. 1:23-cv-08292-SHS

JONATHAN ALTER, *et al.*, on behalf of themselves and all others similarly situated,
 	
   Plaintiffs,
 	
     v.
 	
OPENAI, INC., OPENAI GP, LLC, OPENAI, LLC, OPENAI OPCO LLC, OPENAI GLOBAL LLC, OAI CORPORATION, LLC, OPENAI HOLDINGS, LLC, and MICROSOFT CORPORATION,
 	
   Defendants.
	Case No. 1:23-cv-10211-SHS

FICTION AND NONFICTION PLAINTIFFS' FIRST SET OF REQUESTS FOR
PRODUCTION OF DOCUMENTS TO DEFENDANT MICROSOFT CORPORATION

PROPOUNDING PARTY: Fiction and Nonfiction Plaintiffs

RESPONDING PARTY: Microsoft Corporation

SET NUMBER: One

   Pursuant to Rules 26 and 34 of the Federal Rules of Civil Procedure, the Fiction and

Nonfiction Plaintiffs request that Defendant Microsoft Corporation ("Microsoft") respond to the

following First Set of Requests for Production of Documents ("Requests"). Responses to these

Requests, as well as any responsive documents, shall be delivered to, or made available for

inspection and copying at, the offices of Lieff, Cabraser, Heimann & Bernstein, LLP, 250

Hudson Street, 8th Floor, New York, NY 10013, or at another place mutually agreed upon by the

parties, within thirty (30) days of the date of service or as otherwise mutually agreed by the

parties. Plaintiffs are amenable to an electronic production, subject to agreement by the parties.

In accordance with Rule 34(b), Microsoft shall provide written responses to the following

Requests and shall produce the requested documents as they are kept in the ordinary and usual

course of business or shall organize and label the documents to correspond with the categories in

these Requests. In accordance with Rule 26(e), Microsoft shall supplement or correct its

responses or productions as necessary.

I.   DEFINITIONS

1.   ACTIONS means the above captioned litigation, *Authors Guild et al. v. Open AI*

Inc. et al., No. 1:23-cv-08292-SHS (S.D.N.Y.) (opened Sept. 19, 2023 and amended December

5, 2023), and *Alter et al. v. OpenAI et al.*, No / 1:23-cv-10211-SHS (S.D.N.Y.) (opened Nov. 21,

2023, and amended Dec. 19, 2023).

2.   AMENDED COMPLAINTS means the complaints in this ACTIONS filed at

Fiction Docket No. 40 and Nonfiction Docket No. 26.

3.   CHATGPT means all consumer-facing versions of the chatbot application

OpenAI released in November 2022 (and discussed in paragraphs 78-83 of the Fiction Amended

Complaint (Fiction Docket No. 40)) and any LLM underlying those consumer-facing

applications.

4.   FICTION CLASS WORKS has the same meaning as the term defined in

paragraph 321 of the Fiction Amended Complaint (Fiction Docket No. 40)

5.   NONFICTION CLASS WORKS means the literary works included in the propose class definition in Paragraph 109 of the Nonfiction Amended Complaint (Nonfiction Docket No. 26).

6.   COMMUNICATIONS means the conveyance (in the form of facts, ideas, thoughts, opinions, data, inquiries or otherwise) of information and INCLUDES correspondence, memoranda, reports, presentations, face-to-face conversations, telephone conversations, text messages, instant messages, voice messages, negotiations, agreements, inquiries, understandings, meetings, letters, notes, mail, email, exchanges of recorded information, and postings of any type. COMMUNICATIONS INCLUDES instances where one party disseminates information that the other party receives but does not respond to.

7.   CONCERN means be the subject of, make reference to, comment on, discuss, describe, identify, or contain text or images about the stated topic.

8.   DOCUMENTS means all materials within the scope of Federal Rule of Civil Procedure 34 INCLUDING all writings and recordings, INCLUDING the originals, drafts and all non-identical copies, whether different from the original by reason of any notation made on such copies or otherwise (INCLUDING email and attachments, correspondence, memoranda, notes, diaries, statistics, letters, telegrams, minutes, contracts, reports, studies, checks, statements, tags, labels, invoices, brochures, periodicals, receipts, returns, summaries, pamphlets, books, interoffice and intra-office communications, instant messages, chats, offers, notations of any sort of conversations, working papers, applications, permits, file wrappers, indices, telephone calls, meetings or printouts, teletypes, telefax, invoices, worksheets, and all drafts, alterations, modifications, changes and amendments of any of the foregoing), graphic or aural representations of any kind (INCLUDING photographs, charts, microfiche, microfilm,

- 3 -

videotape, recordings, motion pictures, plans, drawings, surveys), and electronic, mechanical, magnetic, optical, or electric records or representations of any kind (INCLUDING computer files and programs, tapes, cassettes, discs, and recordings), INCLUDING METADATA.

9.   ELECTRONICALLY STORED INFORMATION (ESI) refers to information and DOCUMENTS within the full scope of Federal Rule of Civil Procedure 34—with all METADATA intact—created, manipulated, communicated, stored, and best utilized in electronic or other digital form, and stored on electronic media. Examples of ESI INCLUDE e-mail, messages posted on electronic message boards, forum postings, support tickets, videos, discussion boards, data, source code, websites, Microsoft Word files, Microsoft Excel files, and instant messages.

10.   INCLUDING means including but not limited to.

11.   LARGE LANGUAGE MODEL or LLM has the same meaning as it does in the Fiction Amended Complaint (Fiction Docket No. 40, which discusses and defines that term in paragraphs 56 through 70.

12.   METADATA refers to structured information about an electronic file that is embedded in the file, describing the characteristics, origins, usage, and validity of the electronic file.

13.   MICROSOFT, or YOU means Microsoft Corporation and any of its directors, officers, employees, partners, members, representatives, agents (INCLUDING attorneys, accountants, consultants, investment advisors or bankers), and any other person acting or purporting to act on their behalf, as well as corporate parents, subsidiaries, affiliates, predecessor entities, successor entities, divisions, departments, groups, acquired entities, related entities, or any other entity acting or purporting to act on its behalf.

14.   OPENAI means the entities addressed in paragraphs 29 through 48 of the Fiction Amended Complaint (Fiction Docket No. 40) and any of their directors, officers, employees, partners, members, representatives, agents (INCLUDING attorneys, accountants, consultants, investment advisors or bankers), and any other person acting or purporting to act on their behalf, as well as corporate parents, subsidiaries, affiliates, predecessor entities, successor entities, divisions, departments, groups, acquired entities, related entities, or any other entity acting or purporting to act on their behalf.

15.   PLAINTIFFS means the plaintiffs in the ACTIONS named in the AMENDED COMPLAINTS.

16.   RELATE TO means refer to, respond to, describe, evidence, or constitute, in whole or in part.

II.   RELEVANT TIME PERIOD

The relevant time period is January 1, 2018 through the present ("Relevant Time Period"), unless otherwise specifically indicated, and shall INCLUDE all DOCUMENTS and any other information relating to such period, even though prepared or published outside of the Relevant Time Period. If a DOCUMENT prepared before the Relevant Time Period is necessary for a correct or complete understanding of any DOCUMENT covered by any of these Requests, please provide the earlier DOCUMENT as well. If any DOCUMENT is undated and the date of its preparation cannot be determined, please produce the DOCUMENT if it is otherwise responsive to any Request.

III.   INSTRUCTIONS

1.   The production by one person, party, or entity of a DOCUMENT does not relieve another person, party, or entity from the obligation to produce his, her, or its own copy of that DOCUMENT.

2.   Produce DOCUMENTS not otherwise responsive to these Requests if such DOCUMENTS RELATE TO the DOCUMENTS that are called for by these Requests, or if such DOCUMENTS are attached to DOCUMENTS called for by these Requests.

3.   Produce each DOCUMENT requested herein in its entirety and without deletion, excisions, redaction, or other modification regardless of whether YOU consider the entire document to be relevant or responsive.

4.   If any DOCUMENT is known to have existed but no longer exists, has been destroyed, or is otherwise unavailable, identify the DOCUMENT, the reason for its loss, destruction, or unavailability, the name of each person known or reasonably believed by MICROSOFT to have had possession, custody, or control of the original and any copy thereof (if applicable), and a description of the disposition of each copy of the DOCUMENT.

5.   If no DOCUMENTS responsive to a particular Request exist, state that no responsive DOCUMENTS exist.

6.   If YOU assert that any of the DOCUMENTS and things requested are protected from discovery by attorney-client privilege, the attorney work product doctrine, or any other evidentiary privilege, specify for each DOCUMENT (1) the grounds asserted as the reason for non-production; (2) the date the document was prepared; (3) the identity of the attorney(s) who drafted or received the Document(s) (if attorney-client privilege or attorney work product protection is claimed); (4) the identity of the parties who prepared or received the document; and (5) the nature of the document.

7.   Construe the conjunctions "and" and "or" non-restrictively or non-exclusively if doing so would bring within the scope of these Requests DOCUMENTS that might otherwise be construed to be outside of their scope.

8.   Construe the use of the singular to include the plural; the use of the masculine, feminine, or neuter gender to include the others; and the use of one form of the verb to include the others if doing so would bring within the scope of these Requests DOCUMENTS that might otherwise be construed to be outside of their scope.

RESPONSE TO REQUESTS FOR PRODUCTION

Microsoft's Involvement in OpenAI

REQUEST FOR PRODUCTION NO. 1:

DOCUMENTS sufficient to demonstrate that YOU are an investor in OPENAI, when you became an investor in OPENAI, and YOUR percentage stake in OPENAI.

REQUEST FOR PRODUCTION NO. 2:

DOCUMENTS sufficient to identify the timeline of YOUR investments in OPENAI and the internal DOCUMENTS YOU prepared or relied on in connection with each investment.

REQUEST FOR PRODUCTION NO. 3:

DOCUMENTS and COMMUNICATIONS sufficient to identify DOCUMENTS RELATING TO the allegations in the AMENDED COMPLAINTS that YOU have received or to which YOU have access.

REQUEST FOR PRODUCTION NO. 4:

DOCUMENTS sufficient to identify each person within MICROSOFT who participated in or was consulted CONCERNING the creation, development, design or production of CHATGPT.

REQUEST FOR PRODUCTION NO. 5:

DOCUMENTS sufficient to identify the title and qualifications of each person identified pursuant to Request No. 4 above.

REQUEST FOR PRODUCTION NO. 6:

A copy of the most current curriculums vitae for each person identified in response to Request No. 5 above.

REQUEST FOR PRODUCTION NO. 7:

DOCUMENTS and COMMUNICATIONS sufficient to ascertain YOUR valuation of YOUR investment in OPENAI, at annual intervals, from 2018 to the present.

Already-Gathered Material

REQUEST FOR PRODUCTION NO. 8:

DOCUMENTS that YOU, as of the date of these Requests, have already gathered or collected to submit to any legislative or executive agency, committee, or other governmental entity in the United States that CONCERN or RELATE TO the allegations in the AMENDED COMPLAINTS, INCLUDING the written testimony and responses to questions for the record from Brad Smith, the Vice Chair and President of Microsoft, to the US Senate Judiciary Committee.

REQUEST FOR PRODUCTION NO. 9:

DOCUMENTS that YOU, as of the date these Requests, have already gathered or collected in support of the DOCUMENTS referred to in Request No. 7, or on which the DOCUMENTS referred to in Request No. 7 rely.

Microsoft Statements

REQUEST FOR PRODUCTION NO. 10:

DOCUMENTS sufficient to substantiate the statement made by MICROSOFT CEO Satya Nadella that the "heavy lifting" for OPENAI's LLM "training" was done by MICROSOFT "compute infrastructure," as detailed in paragraph 119 of the AMENDED COMPLAINT.

REQUEST FOR PRODUCTION NO. 11:

DOCUMENTS sufficient to substantiate the statements made by YOUR CEO Satya

Nadella RELATED TO YOUR ownership of "significant rights" and "all IP rights" in OPENAI,

as detailed in paragraphs 53-54 of the AMENDED COMPLAINT.

REQUEST FOR PRODUCTION NO. 12:

DOCUMENTS sufficient to show YOUR interest in any intellectual property developed

or practiced by OPENAI in connection with its LLM products.

REQUEST FOR PRODUCTION NO. 13:

DOCUMENTS RELATED TO the statement made by MICROSOFT CEO Satya Nadella

that Microsoft has "significant rights" in OpenAI and that "We [MICROSOFT] are below them

[OPENAI], above them, around them," as detailed in paragraph 53 of the AMENDED

COMPLAINT.

REQUEST FOR PRODUCTION NO. 14:

DOCUMENTS RELATED TO the statement made by MICROSOFT CEO Satya Nadella

that "We [MICROSOFT] have all the IP rights and all the capability. If OpenAI disappeared

tomorrow, I don't want any customer of ours to be worried about it quite honestly, because we

have all of the rights to continue the innovation. Not just to serve the product, but we can go and

just do what we were doing in partnership ourselves. We have the people, we have the compute,

we have the data, we have everything," as detailed in paragraph 54 of the AMENDED

COMPLAINT.

Plaintiffs

REQUEST FOR PRODUCTION NO. 15:

DOCUMENTS and COMMUNICATIONS CONCERNING or RELATING TO

PLAINTIFFS or CLASS WORKS.

Contributory Infringement

REQUEST FOR PRODUCTION NO. 16:

DOCUMENTS and COMMUNICATIONS CONCERNING or RELATING TO YOUR

right to supervise, direct, and/or control the training of CHATGPT.

REQUEST FOR PRODUCTION NO. 17:

DOCUMENTS and COMMUNICATIONS CONCERNING or RELATING TO the use

of commercial works of fiction to train CHATGPT.

REQUEST FOR PRODUCTION NO. 18:

DOCUMENTS and COMMUNICATIONS CONCERNING or RELATING TO the use

of commercial works of nonfiction to train CHATGPT.

REQUEST FOR PRODUCTION NO. 19:

DOCUMENTS sufficient to determine any manner in which any published works of

fiction or nonfiction, in whole or in part, in any form, format, or language, reside in, are

otherwise stored by CHATGPT.

REQUEST FOR PRODUCTION NO. 20:

DOCUMENTS sufficient to determine any manner in which CHATGPT can reference or

access any published works of fiction or nonfiction, in whole or in part, in any form, format, or

language.

REQUEST FOR PRODUCTION NO. 21:

DOCUMENTS sufficient to determine any manner in which any LLM operated by

OPENAI has referenced or accessed any published works of fiction or nonfiction, in whole or in

part, in any form, format, or language.

REQUEST FOR PRODUCTION NO. 22:

DOCUMENTS and COMMUNICATIONS sufficient to identify the

ELECTRONICALLY STORED INFORMATION through which YOU accessed commercial

works of fiction and/or nonfiction, INCLUDING FICTION CLASS WORKS AND

NONFICTION CLASS WORKS, to train CHATGPT.

REQUEST FOR PRODUCTION NO. 23:

DOCUMENTS and COMMUNICATIONS reflecting or discussing how YOU or

OPENAI accessed any commercial works of fiction or nonfiction, INCLUDING FICTION

CLASS WORKS AND NONFICTION CLASS WORKS, used to train CHATGPT.

REQUEST FOR PRODUCTION NO. 24:

DOCUMENTS sufficient to identify all policies, procedures, or practices YOU or

OPENAI had in place at any time to ensure that training CHATGPT did not infringe on or

otherwise violate another party's intellectual property rights.

REQUEST FOR PRODUCTION NO. 25:

DOCUMENTS and COMMUNICATIONS sufficient to identify any rights clearance

efforts undertaken RELATED TO the material used to train CHATGPT including, but not

limited to, copyright searches and license requests.

REQUEST FOR PRODUCTION NO. 26:

DOCUMENTS and COMMUNICATIONS between YOU and OPENAI related to the

contents of the training dataset used to train CHATGPT.

REQUEST FOR PRODUCTION NO. 27:

DOCUMENTS and COMMUNICATIONS between YOU and OPENAI related to the

process of collecting, maintaining, and using data to train CHATGPT, including the number of

reproductions made of training data in the course of training CHATGPT.

REQUEST FOR PRODUCTION NO. 28:

DOCUMENTS CONCERNING or RELATING TO any complaints YOU have received

regarding YOUR use of potentially copyrighted material in training CHATGPT.

REQUEST FOR PRODUCTION NO. 29:

DOCUMENTS CONCERNING or RELATING TO YOUR policies or procedures

CONCERNING or RELATING TO receiving, processing, or responding to complaints regarding

YOUR use of potentially copyrighted material in training CHATGPT.

REQUEST FOR PRODUCTION NO. 30:

DOCUMENTS and COMMUNICATIONS reflecting or discussing any instances, or

alleged instances, in which CHATGPT directly reproduced potentially copyrighted material.

REQUEST FOR PRODUCTION NO. 31:

DOCUMENTS and COMMUNICATIONS reflecting or discussing your response to any

instances, or alleged instances, in which CHATGPT directly reproduced potentially copyrighted

material.

REQUEST FOR PRODUCTION NO. 32:

DOCUMENTS CONCERNING or RELATING TO any responses YOU have made to

the complaints or instances described in Request Nos. 22 through 24.

REQUEST FOR PRODUCTION NO. 33:

DOCUMENTS and COMMUNICATIONS exchanged between YOU and OPENAI

CONCERNING the use of purportedly copyrighted material in training CHATGPT.

Microsoft's Corporate Structure as Relevant to Issues in the Case

REQUEST FOR PRODUCTION NO. 34:

DOCUMENTS sufficient to identify the aspects of YOUR corporate structure as relevant

to YOUR investment in, involvement in, and monitoring of OPENAI and to the use of

CHATPGPT at MICROSOFT. As to those discrete areas, Plaintiffs request YOUR corporate

structure, INCLUDING, but not limited to, any DOCUMENTS identifying parent companies,

divisions, subsidiaries, affiliates, trade names, fictitious entities, holding companies, related

business entities and/or current directors and officers, as it existed at any time during the

Relevant Time Period.

REQUEST FOR PRODUCTION NO. 35:

DOCUMENTS sufficient to identify YOUR current and former officers, directors,

managers, employees and consultants with knowledge of CHATGPT, business practices

regarding the same, and commercial use of CHATGPT.

REQUEST FOR PRODUCTION NO. 36:

DOCUMENTS sufficient to show the roles and responsibilities of and/or supervisory

relationship between YOUR current and former officers, directors, managers, employees and

consultants and CHATGPT, business practices regarding the same, and commercial use of

regarding the same.

Microsoft's Hiring OpenAI Employees

REQUEST FOR PRODUCTION NO. 37:

DOCUMENTS and COMMUNICATIONS RELATED TO the hiring of OPENAI CEO

Sam Altman on or around November 20, 2023, INCLUDING the roles and responsibilities to

which YOU and Mr. Altman agreed.

REQUEST FOR PRODUCTION NO. 38:

DOCUMENTS and COMMUNICATIONS RELATED TO the hiring of OPENAI

employees other than OPENAI CEO Sam Altman in November 2023, INCLUDING the roles

and responsibilities to which said employees had agreed.

Licensing and Remedy

REQUEST FOR PRODUCTION NO. 39:

DOCUMENTS and COMMUNICATIONS sufficient to identify any efforts to mitigate

and/or remedy instances in which YOU purportedly used copyrighted material without

permission in training CHATGPT.

REQUEST FOR PRODUCTION NO. 40:

DOCUMENTS and COMMUNICATIONS sufficient to show any deals, contracts,

agreements, negotiations, memoranda of understanding, or deal points, whether finalized or not,

whether accepted or not, in executed, draft form, or otherwise, CONCERNING the licensing or

purchase of data, INCLUDING data derived from or about books, to train CHATGPT.

Dated: January 24, 2024   Respectfully submitted,

*/s/ Rachel Geman*_____
Rachel Geman
LIEFF CABRASER HEIMANN & BERNSTEIN, LLP
250 Hudson Street, 8th Floor
New York, NY 10013-1413
Telephone: 212.355.9500
rgeman@lchb.com

Reilly T. Stoler (*pro hac vice forth coming*)
Wesley Dozier (*pro hac vice*)
LIEFF CABRASER HEIMANN & BERNSTEIN, LLP
275 Battery Street, 29th Floor
San Francisco, CA 94111-3339
Telephone: 415.956.1000
rstoler@lchb.com
wdozier@lchb.com

/s/ *Justin Nelson*_____
Justin A. Nelson (*pro hac vice*)
Alejandra C. Salinas (*pro hac vice*)
SUSMAN GODFREY L.L.P.
1000 Louisiana Street, Suite 5100
Houston, TX 77002
Telephone: 713-651-9366
jnelson@susmangodfrey.com
asalinas@susmangodfrey.com

Rohit D. Nath (*pro hac vice*)
SUSMAN GODFREY L.L.P.
1900 Avenue of the Stars, Suite 1400
Los Angeles, California 90067
Telephone: 310-789-3100
rnath@susmangodfrey.com

J. Craig Smyser
SUSMAN GODFREY L.L.P.
1901 Avenue of the Americas, 32nd Floor
New York, New York 10019
Telephone: 212-336-8330
csmyser@susmangodfrey.com

*/s/ Scott Sholder*_____
Scott J. Sholder
CeCe M. Cole
COWAN DEBAETS ABRAHAMS & SHEPPARD LLP
41 Madison Avenue, 38th Floor
New York, New York 10010
Telephone: 212.974.7474
ssholder@cdas.com
ccole@cdas.com

Attorneys for Plaintiffs and the Proposed Classes

PROOF OF SERVICE VIA ELECTRONIC MAIL

On January 24, 2024, per the parties' agreements, I directed the following document to be

served via electronic mail:

**FICTION AND NONFICTION PLAINTIFFS' FIRST SET OF REQUESTS FOR
PRODUCTION OF DOCUMENTS TO DEFENDANT MICROSOFT
CORPORATION**

upon: ahurst@orrick.com; ccariello@orrick.com, mrshapiro@orrick.com,

openaicopyrightlitigation.lwteam@lw.com, and OpenAICopyright@mofo.com

/s/ Rachel Geman
Rachel Geman

- 16 -

UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

AUTHORS GUILD, et al.,	**ECF CASE**
Plaintiffs,	No. 1:23-cv-08292-SHS;
v.	No. 1:23-cv-10211-SHS
OPEN AI INC., et al.,	**PLAINTIFFS' SECOND SET OF**
Defendants.	**REQUESTS FOR PRODUCTION**
	TO MICROSOFT
	CORPORATION
JONATHAN ALTER, et al.,	
Plaintiffs,	
v.	
OPENAI, INC., et al.,	
Defendants.	

PLAINTIFFS' SECOND SET OF REQUESTS FOR PRODUCTION TO DEFENDANT

MICROSOFT CORPORATION

Pursuant to Rules 26 and 34 of the Federal Rules of Civil Procedure, Plaintiffs request that

Defendant Microsoft Corporation ("Microsoft") respond to the following Second Set of Requests

for Production of Documents ("Requests"). Responses to these Requests shall be due within thirty

(30) days of the date of service or as otherwise mutually agreed by the parties. Plaintiffs are

amenable to an electronic production, subject to agreement by the parties.

In accordance with Rule 34(b), Microsoft shall provide written responses to the following

Requests and shall produce the requested documents as they are kept in the ordinary and usual

course of business or shall organize and label the documents to correspond with the categories in

these Requests. In accordance with Rule 26(e), Microsoft shall supplement or correct its responses or productions as necessary.

I.   **DEFINITIONS**

1.   **"Actions"** means the above captioned litigation, *Authors Guild et al. v. Open AI Inc. et al.*, No. 1:23-cv-08292-SHS (S.D.N.Y.) (opened Sept. 19, 2023 and amended December 5, 2023) ("Fiction Docket"), and *Alter et al. v. OpenAI et al.*, No / 1:23-cv-10211-SHS (S.D.N.Y.) (opened Nov. 21, 2023, and amended Dec. 19, 2023) ("Nonfiction Docket").

2.   **"Class Works"** pertains to both Fiction Class Works and Nonfiction Class Works as defined in ¶394 and in ¶397 of the First Consolidated Complaint (Fiction Docket No. 69).

3.   **"Communication(s)"** means the transmittal of information (in the form of facts, ideas, inquiries or otherwise) by any means, including, but not limited to, telephone calls, emails (whether via company server or personal webmail or similar accounts), faxes, text messages (on work or personal phones), instant messages, Skype, Line, WhatsApp, WeChat, other electronic messages, letters, notes, and voicemails.

4.   **"Document(s)"** is defined to be synonymous in meaning and equal in scope to the usage of the term "documents or electronically stored information" in Rule 34(a)(1)(A). For the avoidance of doubt, **Document(s)** includes **Communication(s)**.

5.   **"You", "Your"**, and **"Microsoft"** means Microsoft Corporation and any of its directors, officers, employees, partners, members, representatives, agents (INCLUDING attorneys, accountants, consultants, investment advisors or bankers), and any other person acting or purporting to act on their behalf, as well as corporate parents, subsidiaries, affiliates, predecessor entities, successor entities, divisions, departments, groups, acquired entities, related entities, or any other entity acting or purporting to act on its behalf.

6.   **"OpenAI"** means OpenAI, Inc., OpenAI GP, LLC, OpenAI, LLC, OpenAI OPCO LLC, OpenAI Global LLC, OAI Corporation, LLC, OpenAI Holdings, and any of their directors, officers, employees, partners, members, representatives, agents (including attorneys, accountants, consultants, investment advisors or bankers), and any other person acting or purporting to act on their behalf, as well as corporate parents, subsidiaries, affiliates, predecessor entities, successor entities, divisions, departments, groups, acquired entities, related entities, or any other entity acting or purporting to act on their behalf.

7.   **"Large Language Model," "LLM," "AI Model(s)," "Generative AI system(s)," "model(s),"** and **"API Product(s)"** have the same meaning as they are used in OpenAI's letter to the Register of Copyrights and Director of the U.S. Copyright Office dated October 30, 2023, "Re: Notice of Inquiry and Request for Comment [Docket No. 2023-06]" and include all models listed or described in https://platform.openai.com/docs/models.

8.   **"Person(s)"** means any individual or entity.

9.   **"Fine Tune(d)," "Fine Tuning," "Pre-Train(ed)," "Pre-Training," "Train[ed],"** and **"Training,"** have the same meaning as the term is discussed in **Your** website materials and statements. See e.g., https://platform.openai.com/docs/guides/fine-tuning ("OpenAI's text generation models have been pre-trained on a vast amount of text. To use the models effectively, we include instructions and sometimes several examples in a prompt. Using demonstrations to show how to perform a task is often called "few-shot learning." Fine-tuning improves on few-shot learning by training on many more examples than can fit in the prompt, letting you achieve better results on a wide number of tasks. Once a model has been fine-tuned, you won't need to provide as many examples in the prompt."); https://openai.com/index/language-unsupervised/ ("These results provide a convincing example that pairing supervised learning methods with unsupervised

pre-training works very well; this is an idea that many have explored in the past, and we hope our result motivates further research into applying this idea on larger and more diverse datasets."); and https://openai.com/index/how-should-ai-systems-behave/ ("The two main steps involved in building ChatGPT work as follows . . . First, we "pre-train" models by having them predict what comes next in a big dataset . . . Then, we "fine-tune" these models on a more narrow dataset that we carefully generate with human reviewers who follow guidelines that we provide them."); https://openai.com/index/gpt-4-research/ ("Interestingly, the base pre-trained model is highly calibrated . . . Note that the model's capabilities seem to come primarily from the pre-training process . . . "); https://openai.com/index/text-and-code-embeddings-by-contrastive-pre-training/ ("In this work, we show that contrastive pre-training on unsupervised data at scale leads to high quality vector representations of text and code."); https://openai.com/index/vpt/ ("We trained a neural network to play Minecraft by Video PreTraining (VPT) on a massive unlabeled video dataset of human Minecraft play, while using only a small amount of labeled contractor data. With fine-tuning, our model can learn to craft diamond tools . . .").

10.   **"ChatGPT"** means all consumer-facing versions of the chatbot application OpenAI released in November 2022 (and discussed in paragraphs 83-95 of the Fist Consolidated Complaint (Fiction Docket No. 69)) and any LLM underlying those consumer-facing applications.

11.   **"Concern"** means be the subject of, make reference to, comment on, discuss, describe, identify, or contain text or images about the stated topic.

12.   **"Including"** means including but not limited to.

13.   **"Investors"** means any person or entity that financially invested in the operations and/or entitie(s) controlled and/or overseen by one or more of the OpenAI defendant entities.

14.   **"Relate to"** means refer to, respond to, describe, evidence, or constitute, in whole or in part.

15.   **"Publishers**" means a person or entity whose business is publishing materials, including, but not limited to, books, periodicals, magazines, and newspapers.

16.   **"Training Material"** means any data (e.g. text) obtained for or used to train a Large Language Model.

II.   RELEVANT TIME PERIOD

The relevant time period is January 1, 2015 through the present ("Relevant Time Period"), unless otherwise specifically indicated, and shall include all **Documents** and any other information relating to such period, even though prepared or published outside of the Relevant Time Period. If a **Document** prepared before the Relevant Time Period is necessary for a correct or complete understanding of any **Document** covered by any of these Requests, please provide the earlier **Document** as well. If any **Document** is undated and the date of its preparation cannot be determined, please produce the **Document** if it is otherwise responsive to any Request.

III.   INSTRUCTIONS

1.   The production by one person, party, or entity of a **Document** does not relieve another person, party, or entity from the obligation to produce his, her, or its own copy of that **Document**.

2.   Produce **Documents** not otherwise responsive to these Requests if such **Documents** relate to the **Documents** that are called for by these Requests, or if such **Documents** are attached to **Documents** called for by these Requests.

3.   Produce each **Document** requested herein in its entirety and without deletion, excisions, redaction, or other modification regardless of whether YOU consider the entire document to be relevant or responsive.

4.   If any **Document** is known to have existed but no longer exists, has been destroyed, or is otherwise unavailable, identify the **Document**, the reason for its loss, destruction, or unavailability, the name of each person known or reasonably believed by Microsoft to have had possession, custody, or control of the original and any copy thereof (if applicable), and a description of the disposition of each copy of the **Document**.

5.   If no **Documents** responsive to a particular Request exist, state that no responsive **Documents** exist.

6.   If You assert that any of the **Documents** and things requested are protected from discovery by attorney-client privilege, the attorney work product doctrine, or any other evidentiary privilege, specify for each **Document** (1) the grounds asserted as the reason for non-production; (2) the date the **Document** was prepared; (3) the identity of the attorney(s) who drafted or received the **Document(s)** (if attorney-client privilege or attorney work product protection is claimed); (4) the identity of the parties who prepared or received the **Document**; and (5) the nature of the **Document**.

7.   Construe the conjunctions "and" and "or" non-restrictively or non-exclusively if doing so would bring within the scope of these Requests **Documents** that might otherwise be construed to be outside of their scope.

8.   Construe the use of the singular to include the plural; the use of the masculine, feminine, or neuter gender to include the others; and the use of one form of the verb to include the others if doing so would bring within the scope of these Requests **Documents** that might otherwise be construed to be outside of their scope.

Plaintiffs, by and through their undersigned attorneys, request that Defendants provide a response to these Requests within thirty (30) days of the date of service hereof as provided by Federal Rule of Civil Procedure 33.

<div align="center">

REQUESTS FOR PRODUCTION

</div>

REQUEST FOR PRODUCTION NO. 41:

Documents sufficient to identify all datasets containing commercial works of fiction and/or nonfiction, including Class Works, that **You** or **OpenAI** accessed, downloaded or copied to train **Large Language Models**.

REQUEST FOR PRODUCTION NO. 42:

Documents Concerning or **Relating to** any complaints **You** have received regarding **Your** or **OpenAI's** use of potentially copyrighted material in training **Large Language Models**.

REQUEST FOR PRODUCTION NO. 43:

Documents Concerning or **Relating to Your** policies or procedures **Relating to** receiving, processing, or responding to complaints regarding **Your** or **OpenAI's** use of potentially copyrighted material in training **Large Language Models**.

REQUEST FOR PRODUCTION NO. 44:

Documents sufficient to identify any effort by **You** or **OpenAI** to remedy or mitigate **Your** or **OpenAI's** unauthorized use of copyrighted material to train any **Large Language Models**.

REQUEST FOR PRODUCTION NO. 45:

All **Communications** between January 1, 2016 and December 31, 2022, related to the use or reproduction of copyrighted materials to train any artificial intelligence model **OpenAI** developed, including without limitation all **Large Language Models** and diffusion models (e.g., DALL-E, Codex, GPT-2).

REQUEST FOR PRODUCTION NO. 46:

Documents sufficient to show all attempts by **You** or **OpenAI** to obtain a license for the use of text or audio works in connection with the development of any **Large Language Model(s),** including but not limited to **Documents** reflecting any negotiations of such licenses.

REQUEST FOR PRODUCTION NO. 47:

All **Documents Concerning** or **Relating to Your** decision(s) to seek or obtain a license to use any copyrighted material as **Training Material**.

REQUEST FOR PRODUCTION NO. 48:

All **Documents Concerning** or **Relating to** the relative value of different types of **Training Material**.

REQUEST FOR PRODUCTION NO. 49:

All **Documents** concerning or relating to **OpenAI** deriving **Training Material** from any physical books or audiobooks.

REQUEST FOR PRODUCTION NO. 50:

All **Documents Concerning or Relating to** OpenAI's use of digital books as **Training Material**.

REQUEST FOR PRODUCTION NO. 51:

Documents sufficient to show any reproduction of the data **OpenAI** used to train any **Large Language Model**, including any reproductions that came into Microsoft's possession.

REQUEST FOR PRODUCTION NO. 52:

All **Documents** and **Communications Concerning** or **Relating** to the use or contemplated use of material from any text data repository (e.g. Library Genesis, Internet Archive, Z-Library,

Common Crawl, WebText, Project Gutenberg, Anna's Archive, Open Library, Reddit, DuXiu) as **Training Material**.

REQUEST FOR PRODUCTION NO. 53:

All **Documents Concerning** or **Relating to OpenAI's** potential exhaustion of available text data for use as **Training Material.**

REQUEST FOR PRODUCTION NO. 54:

Documents sufficient to show when **You** became aware of each dataset used to train **OpenAI's Large Language Models**, including but not limited to, all **Documents Concerning** and **Relating to** the "writeup of OpenAI's technical results on large language models**"** **OpenAI** sent to **You** in February 2020 that "include[d] details on training methods, datasets, etc to facilitate comparisons and collaboration." (MSFT_AICPY_000228653).

REQUEST FOR PRODUCTION NO. 55:

All **Documents Concerning** or **Relating to** the relative value of different types of text data in training **Large Language Models**.

REQUEST FOR PRODUCTION NO. 56:

All **Documents Concerning** or **Relating to** any programming of any **Large Language Models** developed by **OpenAI** designed or intended—in whole or in part—to avoid copyright infringement and/or to avoid reproducing materials used to train the **Large Language Models** or **ChatGPT**.

REQUEST FOR PRODUCTION NO. 57:

All **Documents Concerning** or **Relating to** the impact that including or excluding human-generated text from a **Large Language Model's** training data has on the quality of its outputs.

REQUEST FOR PRODUCTION NO. 58:

Documents sufficient to show any exclusive or non-exclusive rights granted to **You** by **OpenAI** to commercialize, monetize, or productize **OpenAI's Large Language Models**, training data, or intellectual property.

REQUEST FOR PRODUCTION NO. 59:

Documents sufficient to identify all of Your commercial products that use or include **any Large Language Model developed by OpenAI** , including but not limited to, **Your** Bing search engine (as set forth in the First Consolidated Complaint [Fiction Docket No. 69] at ¶166) as well as Microsoft Teams, Microsoft 365 Copilot and GitHub Copilot (*id*. at ¶167).

REQUEST FOR PRODUCTION NO. 60:

Documents sufficient to identify each person within **Microsoft** who participated, was involved in or was consulted **Concerning** the use, implementation or incorporation of **ChatGPT** or **Large Language Models** in **Your** commercial products.

REQUEST FOR PRODUCTION NO. 61:

All **Documents Concerning** or **Relating to** customers substituting away from or towards a **Microsoft** product using a **Large Language Model** based on the quality of the **Large Language Model**.

REQUEST FOR PRODUCTION NO. 62:

Documents sufficient to identify **Your** gross revenues, net revenues, and profits, by month, from Your commercial products that include or use **ChatGPT** or **Large Language Models**.

REQUEST FOR PRODUCTION NO. 63:

Documents sufficient to show the gross revenues, net revenues, and profits, by month, generated by each **Large Language Model You** have commercialized, sold, and/or licensed, either as a standalone product or as part of another product.

REQUEST FOR PRODUCTION NO. 64:

Document sufficient to show Your balance sheet, income statement, and cash flow statement, on a monthly basis during the relevant time period.

REQUEST FOR PRODUCTION NO. 65:

All **Documents Concerning** or **Relating to** predictions, forecasts, or projections of profits, revenues or cash flows of **OpenAI** or from **Large Language Models** or products using or containing **Large Language Models**.

REQUEST FOR PRODUCTION NO. 66:

All presentations, memoranda, or other **Documents** provided to **Your** Board of Directors **Concerning** or **Relating to OpenAI**'s development of Large Language Models.

REQUEST FOR PRODUCTION NO. 67:

Documents sufficient to show each version of the terms and/or conditions of use of **Your** products that use or include **ChatGPT** or **Large Language Models**.

REQUEST FOR PRODUCTION NO. 68:

Retention agreements You entered from January 1, 2016 to the present, with any law firm(s) for actual or potential copyright litigation **Related to Your** investment in and/or partnership with **OpenAI**.

REQUEST FOR PRODUCTION NO. 69:

Documents sufficient to show all indemnification provisions agreed to by and between **You** and **OpenAI**.

REQUEST FOR PRODUCTION NO. 70:

All **Documents Concerning** or **Relating to** whether **OpenAI's** use of copyrighted material as Training Material complies with copyright law in the United States or elsewhere.

REQUEST FOR PRODUCTION NO. 71:

All **Documents You** received from OpenAI in the course of **Your** due diligence regarding each of **Your** investments in **OpenAI**.

REQUEST FOR PRODUCTION NO. 72:

Documents and **Communications Concerning** or **Relating to Your** decision to modify **Your** normal due diligence process for the purposes of **Your** investment in **OpenAI**.

REQUEST FOR PRODUCTION NO. 73:

Documents sufficient to identify all data that was shared among **You** and **OpenAI** in the course of **Your** partnership with **OpenAI**, including but not limited to data shared pursuant to the terms of the Joint Development and Collaboration Agreement (the "JDCA") (initial JDCA produced at MSFT_AICPY_000000345) since June 2019.

REQUEST FOR PRODUCTION NO. 74:

Documents sufficient to identify each of **Your** employees who at any point had access to "Restricted Access Model Training Code" pursuant to Section 3(e) of the Second Amended and Restated JDCA (the "Second Amended JDCA") (MSFT_AICPY_000000575).

REQUEST FOR PRODUCTION NO. 75:

Documents sufficient to identify any "secure repository" **You** identified to **OpenAI** pursuant to Section 3(e)(iv) the Amended and Restated JDCA (MSFT_AICPY_000004210).

REQUEST FOR PRODUCTION NO. 76:

Documents sufficient to identify all individuals involved in copying the "Development Materials"—as defined in the Second Amended JDCA—for storage in **Your** secure repository, as set forth in Section 3(h)(iv) the Second Amended JDCA.

REQUEST FOR PRODUCTION NO. 77:

All **Documents Concerning** or **Relating to** the Data Working Group, as referenced in the Second Amended JDCA, including but not limited to all **Communications** among members of, meeting minutes of, and data policies compiled by the Data Working Group.

REQUEST FOR PRODUCTION NO. 78:

All **Documents** created in the course of the Governing Board's—as referenced in the Second Amended JDCA—duties, including but not limited to meeting minutes, memorandum and **Communications**.

Dated: August 14, 2024   Respectfully submitted,

/s/ Rachel Geman_____
Rachel Geman
Wesley Dozier (*pro hac vice*)
Anna J. Freymann
LIEFF CABRASER HEIMANN & BERNSTEIN, LLP
250 Hudson Street, 8th Floor
New York, NY 10013-1413
Telephone: 212.355.9500
rgeman@lchb.com
wdozier@lchb.com
afreymann@lchb.com

Reilly T. Stoler (*pro hac vice*)
LIEFF CABRASER HEIMANN & BERNSTEIN, LLP
275 Battery Street, 29th Floor
San Francisco, CA 94111-3339
Telephone: 415.956.1000
rstoler@lchb.com

/s/ *Justin Nelson*   _
Justin A. Nelson (*pro hac vice*)
Alejandra C. Salinas (*pro hac vice*)
SUSMAN GODFREY L.L.P.
1000 Louisiana Street, Suite 5100
Houston, TX 77002
Telephone: 713-651-9366
jnelson@susmangodfrey.com
asalinas@susmangodfrey.com

Rohit D. Nath (*pro hac vice*)
SUSMAN GODFREY L.L.P.
1900 Avenue of the Stars, Suite 1400
Los Angeles, California 90067
Telephone: 310-789-3100
rnath@susmangodfrey.com

J. Craig Smyser
SUSMAN GODFREY L.L.P.
1901 Avenue of the Americas, 32nd Floor
New York, New York 10019
Telephone: 212-336-8330
csmyser@susmangodfrey.com

/s/ Scott Sholder
Scott J. Sholder
CeCe M. Cole
COWAN DEBAETS ABRAHAMS & SHEPPARD LLP
41 Madison Avenue, 38th Floor
New York, New York 10010
Telephone: 212.974.7474
ssholder@cdas.com
ccole@cdas.com

Attorneys for Plaintiffs and the Proposed Classes

<u>CERTIFICATE OF SERVICE</u>

I hereby certify that on August 14, 2024, a copy of the foregoing was served via electronic

mail to all counsel of record in this matter.

<div align="right">

/s/ Ariana Delucchi
(Signature)

</div>

3079344.1